ZooKeys 1124: 1–21 (2022)
doi: 10.3897/zookeys.1124.83811
https://zookeys.pensoft.net

RESEARCH ARTICLE

A peer-reviewed open-access journal
ZooKeys
Launched to accelerate biodiversity research

Novel gene rearrangement pattern in mitochondrial genome of *Ooencyrtus plautus* Huang & Noyes, 1994: new gene order in Encyrtidae (Hymenoptera, Chalcidoidea)

Zhi-Ping Xing[1,2], Xin Liang[1,2], Xu Wang[3,4], Hao-Yuan Hu[1,2], Yi-Xin Huang[1,2,3]

1 *Collaborative Innovation Center of Recovery and Reconstruction of Degraded Ecosystem in Wanjiang Basin Co-founded by Anhui Province and Ministry of Education, Wuhu, Anhui 241000, China* **2** *School of Ecology and Environment, Anhui Normal University, Wuhu, Anhui 241000, China* **3** *Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, 1 Beichen West Road, Chaoyang District, Beijing, 100101, China* **4** *Anhui Provincial Key Laboratory of the Conservation and Exploitation of Biological Resources, College of Life Sciences, Anhui Normal University, Wuhu, Anhui 241000, China*

Corresponding author: Hao-Yuan Hu (haoyuanhu@126.com), Yi-Xin Huang (huangyx@ahnu.edu.cn)

Academic editor: Francisco J. P. Felipo | Received 15 March 2022 | Accepted 14 September 2022 | Published 10 October 2022

https://zoobank.org/64044D13-AFB2-439F-BAFA-4FEA816D78EA

Citation: Xing Z-P, Liang X, Wang X, Hu H-Y, Huang Y-X (2022) Novel gene rearrangement pattern in mitochondrial genome of *Ooencyrtus plautus* Huang & Noyes, 1994: new gene order in Encyrtidae (Hymenoptera, Chalcidoidea). ZooKeys 1124: 1–21. https://doi.org/10.3897/zookeys.1124.83811

Abstract

Studies of mitochondrial genomes have a wide range of applications in phylogeny, population genetics, and evolutionary biology. In this study, we sequenced and analyzed the mitochondrial genome of *Ooencyrtus plautus* Huang & Noyes, 1994 (Hymenoptera, Encyrtidae). The nearly complete mitogenome of *O. plautus* was 15,730 bp in size, including 13 PCGs (protein-coding genes), 22 tRNAs, 2 rRNAs, and a nearly complete control region. The nucleotide composition was significantly biased toward adenine and thymine, with an A + T content of 84.6%. We used the reference sequence of *Chouioia cunea* and calculated the Ka/Ks ratio for each set of PCGs. The highest value of the Ka/Ks ratio within 13 PCGs was found in *nad2* with 1.1, suggesting that they were subjected to positive selection. This phenomenon was first discovered in Encyrtidae. Compared with other encyrtid mitogenomes, a translocation of *trnW* was found in *O. plautus*, which was the first of its kind to be reported in Encyrtidae. Comparing with ancestral arrangement pattern, wasps reflect extensive gene rearrangements. Although these insects have a high frequency of gene rearrangement, species from the same family and genus tend to have similar gene

Copyright *Zhi-Ping Xing et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

sequences. As the number of sequenced mitochondrial genomes in Chalcidoidea increases, we summarize some of the rules of gene rearrangement in Chalcidoidea, that is four gene clusters with frequent gene rearrangements. Ten mitogenomes were included to reconstruct the phylogenetic trees of Encyrtidae based on both 13 PCGs (nucleotides of protein coding genes) and AA matrix (amino acids of protein coding genes) using the maximum likelihood and Bayesian inference methods. The phylogenetic tree reconstructed by Bayesian inference based on AA data set showed that *Aenasius arizonensis* and *Metaphycus eriococci* formed a clade representing Tetracneminae. The remaining six species formed a monophyletic clade representing Encyrtinae. In Encyrtinae, *Encyrtus* forms a monophyletic clade as a sister group to the clade formed by *O. plautus* and *Diaphorencyrtus aligarhensis*. *Encyrtus sasakii* and *Encyrtus rhodooccisiae* were most closely related species in this monophyletic clade. In addition, gene rearrangements can provide a valuable information for molecular phylogenetic reconstruction. These results enhance our understanding of phylogenetic relationships among Encyrtidae.

Keywords
Encyrtinae, mitogenome, pairwise breakpoint distance, phylogenetic tree, Tetracneminae

Introduction

The mitochondrial genome is a standard circular molecule, mostly range between 15 kb and 18 kb in size, with 37 genes, including 13 protein-coding genes, three of which are oxidative phosphorylation complexes, 2 rRNAs, and 22 tRNAs and a major non-coding region which mainly regulates replication and transcription (Fig. 1) (Boore 1999; Cameron 2014). The mitochondrial (mt) genome has the characteristics of gene recombination, maternal inheritance, conservation of gene components, and high AT content, and it is considered to be an ideal molecular marker for species identification and phylogenetic or evolutionary studies (Boore and Brown 1998; Curole and Kocher 1999). In addition, the gene content, genome size, and RNA secondary structure of the mt genome can also provide useful information for phylogeny (Boore and Brown 1998; Gissi et al. 2008; Cameron 2014). Gene rearrangement is one of the most frequently studied features in animal mitochondrial genomes (Boore and Brown 1998; Cameron 2014; Li et al. 2016). It is usually conserved in major lineages but may be rearranged in some groups (Boore and Brown 1998; Boore 1999; Dowton et al. 2002a). The rearrangement processes of mitochondrial genes can be described as transposition, inversion, reverse transposition, and TDRL (tandem-duplication-random-loss) (Dowton et al. 2002b; Cameron 2014). The movement of TDRL is to describe the duplication of multiple consecutive genes and the successive random loss of one of the two copies (Bernt et al. 2007). Previous works by Sankoff et al. (1992) and Boore et al. (1998) showed that the gene sequence of the mitochondrial genome contains phylogenetic signals. Studying gene rearrangements in lower taxonomic lineages of insects can provide some evidence for the evolution of these groups (Mao et al. 2014; Li et al. 2016).

Extensive mitochondrial genome data indicates that, compared to other orders in the hexapoda, Hymenoptera (Dowton and Austin 1999; Dowton et al. 2009; Wei et al. 2014) and Hemiptera (Shao and Barker 2003; Shao et al. 2003; Johnson et al. 2004)

have highly accelerated and independent gene arrangement events. Chalcidoidea is one of the most diverse groups of insects in the order Hymenoptera. Parasitic chalcidoids are widely used as biological control of various agricultural pests (Heraty 2017). At the same time, diverse morphology, body sizes, lifestyles, and different types of parasitism, which is also reflected in their amazing genetic arrangement (Austin et al. 1998; Heraty et al. 2013). The mitochondrial genome of Chalcidoidea has been confirmed to have a large number of gene rearrangements (Oliveira et al. 2008; Xiao et al. 2011). A total of 40 species of mitochondrial genomes have been reported in the Chalcidoidea, including 11 families and all sequenced mitochondrial genomes have rearrangements compared with putative ancestor (Fig. 2) (up to date: 15 March 2022). The parasitic life history, body size, and developmental period in Hymenoptera are hypothesized to be related to their gene rearrangements (Dowton and Austin 1995; Shao et al. 2001), and the information of these gene rearrangements may be valuable for the phylogenetic reconstruction of specific lineages (Yuan et al. 2015; Liu et al. 2017).

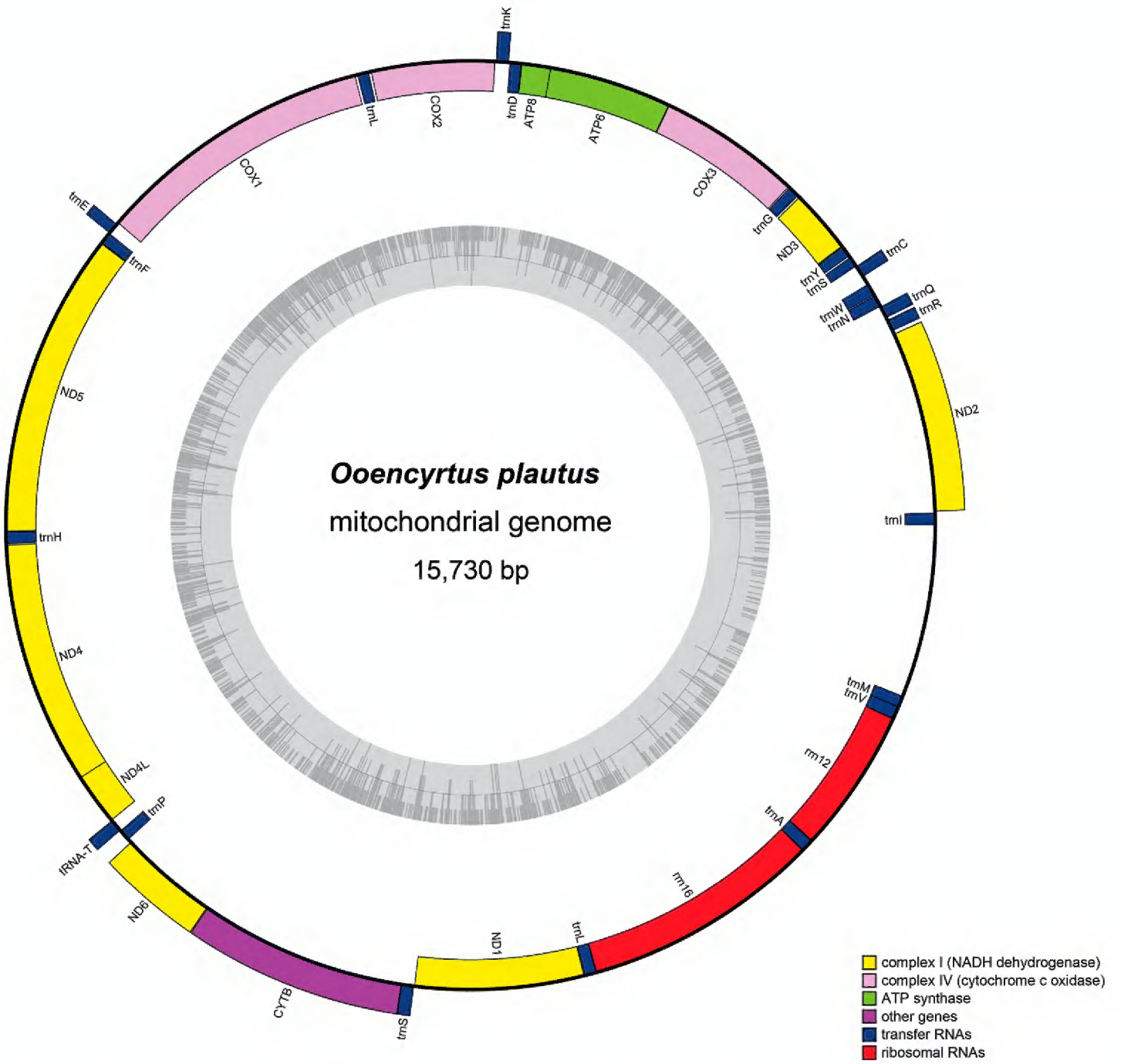

Figure 1. Circular map of the *Ooencyrtus plautus* mitochondrial genome.

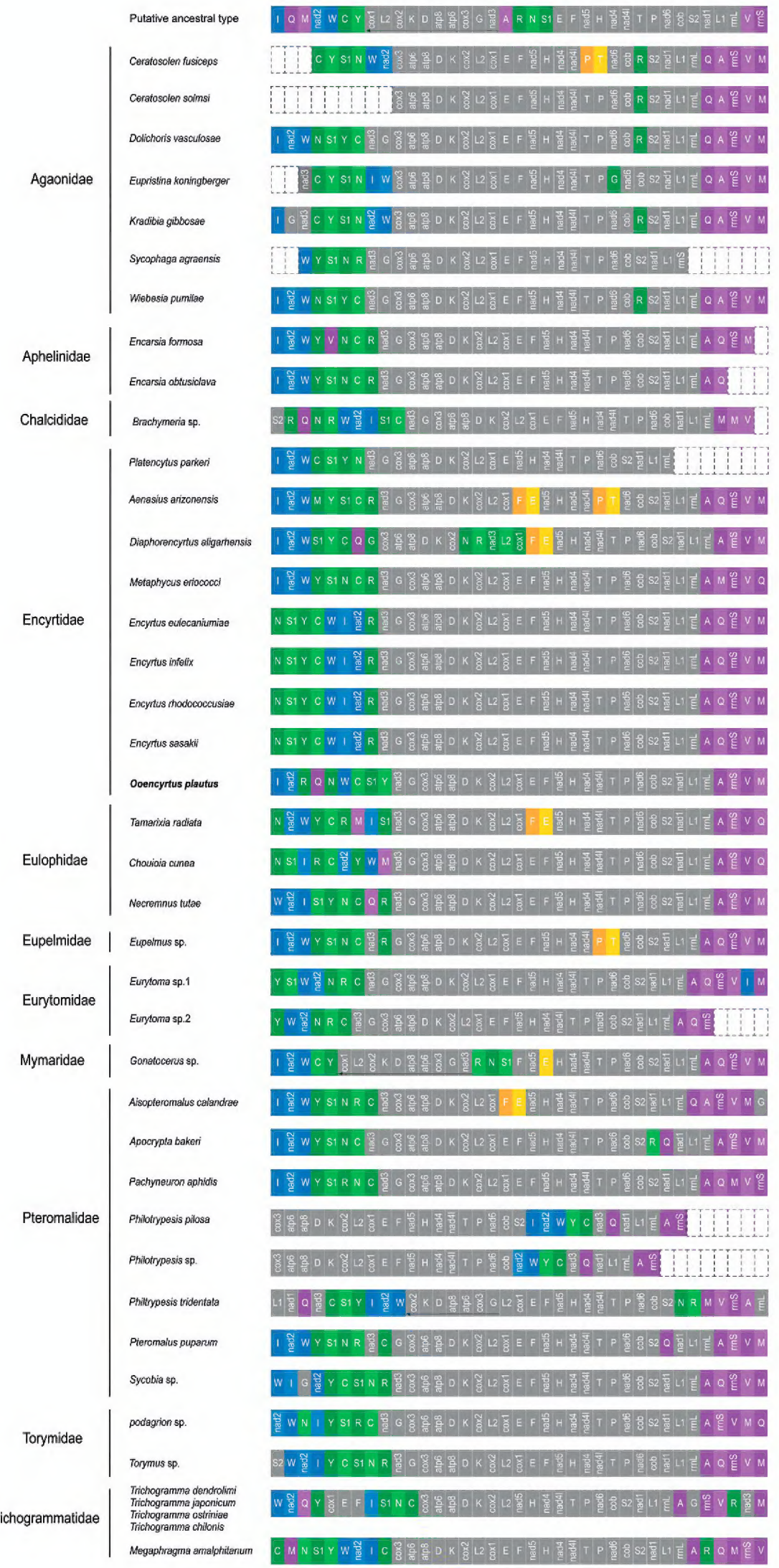

Figure 2. Mitochondrial genome organization in Chalcidoidea. Colored boxes indicate rearranged gene clusters, and gray boxes indicate conserved gene blocks.

Ooencyrtus plautus is a kind of parasitoid in the family Encyrtidae, which has the characteristics of high parasitism rate and strong reproduction ability. In the Encyrtidae, seven species of complete mitochondrial genes and one species of partial mitochondrial genes have been reported (Table 1). In this study, the nearly complete mitochondrial genome of *O. plautus* was measured, sequenced, and annotated. The characteristics of the mitogenomes are described in terms of genome structure, nucleotide content, frequency of usage of start and stop codons, and gene rearrangement. It reported a new gene rearrangement which was found in the mitochondrial genome of *O. plautus* (The long-distance transposition of *trnW* and the reverse transposition of gene cluster *trnN-trnS1-trnY*) and conducted phylogenetic analyses using the current mitogenome along with those previously reported from Encyrtidae.

Table 1. List of species investigated and their related information.

	Family	Subfamily	Taxa	GenBank Accession No.	Location/Refence
1	Aphelinidae		*Encarsia formosa*	MG813797.1	(Zhu et al. 2018)
2			*Encarsia obtusiclava*	MG813798.1	(Zhu et al. 2018)
3	Encyrtidae	Tetracneminae	*Aenasius arizonensis*	MK630013	(Ma et al. 2019)
4		Tetracneminae	*Metaphycus eriococci*	MW255970	Direct Submission
5		Encyrtinae	*Diaphorencyrtus aligarhensis*	MN274569	(Du et al. 2019)
6		Encyrtinae	*Encyrtus eulecaniumiae*	NC_051459	Direct Submission
7		Encyrtinae	*Encyrtus infelix*	MH729198	(Xiong et al. 2019)
8		Encyrtinae	*Encytus rhodooccisiae*	NC_051460	Direct Submission
9		Encyrtinae	*Encyrtus sasakii*	NC_051458	Direct Submission
10		Encyrtinae	*Ooencyrtus plautus*	OP442361	This study

Materials and methods

Sample collection and DNA extraction

Specimens of *Ooencyrtus plautus* were collected from Fuzhou city, Fujian province, in September 2020. They were reared in the laboratory, then processed for DNA extraction. Total genomic DNA was extracted using the cetyltrimethyl ammonium bromide (CTAB) method (Shahjahan et al. 1995).

Sequencing and assembly

Sequencing was performed using a whole genome shotgun (WGS) strategy on the Illumina Novaseq platform. The quality of data was checked using FastQC (Andrews 2013). Raw data were filtered into high-quality data after processing with AdapterRemoval v. 2. Then, Single Nucleotide Polymorphism (SNP), Insertion and Deletion (InDel), Copy Number Variation (CNV), and Structural Variation (SV) analysis were conducted to ensure the reliability of nucleotides. The assembly of the mitochondrial genome was accomplished with Novoplasy v. 2.7 (Dierckxsens et al. 2017).

Mitochondrial genome annotation

Gene annotation of mitochondrial genome was performed using MitoZ and 13 protein-coding genes as well as 22 tRNA genes are annotated (Meng et al. 2019). Twenty-two tRNA genes were verified with the use of MITOS WebServer, setting the parameters with the Invertebrate Mito genetic code (Minh et al. 2013). Every sequence of tRNA genes was manually checked separately. According to the secondary structure of tRNA predicted by MITOS, it was drawn manually using Adobe Illustrator. Annotation of the 13 protein-coding genes were refined manually by identifying the corresponding open reading frames using the invertebrate mitochondrial code. The mitogenome maps were produced using Organellar Genome DRAW (OGDRAW) (Marc et al. 2013).

Comparative analysis

Geneious v. 11.0.2 was used to examine all genes in the mitochondrial genome (Kearse et al. 2012). MEGA X (Sudhir et al. 2018) was used to analyze base composition and relative synonymous codon usage (RSCU). To calculate the number of synonymous substitutions for each synonymous site (Ks) and the number of non-synonymous substitutions for each non-synonymous site (Ka) through DnaSP 5. AT/GC skewness is calculated as AT skewness = (A – T) / (A + T) and GC-skew = (G – C) / (G + C) (Rozas et al. 2003). To measure the relative composition of different bases was measured based on GC and AT skew.

Phylogenetic analysis

In this study, a total of 10 species of mitogenomes were analyzed, of which *Encarsia formosa* and *Encarsia obtusiclava* of family Aphelinidae were selected as outgroups (Table 1). MAFFT v. 7.3.1 was used to align the PCGs (Kazutaka and Standley 2016). Alignments of individual genes were concatenated to generate two kind of data sets: 1) the PCG matrix, including all three codon positions of protein-coding genes; 2) the AA matrix, translated 13 PCG into amino acids.

Both ML (maximum likelihood) and BI (Bayesian inference) analyses were performed on the concatenated data set used for phylogenetic reconstruction. In W-IQtree (Trifinopoulos et al. 2016), the best-fit model was used for maximum likelihood analysis. The site-heterogeneous mixture model (CAR+GTR) was used in PhyloBayes analysis. The trees were sampled every 1000 generations (Huelsenbeck 2012). FigTree v. 1.3.1 was used to view the generated tree.

Results

Genome structure and organization

The nearly complete mitogenome of *O. plautus* was 15,730 bp in size, including 13 PCGs, 22 tRNAs, 2 rRNAs, and a nearly complete control region. In *Ooencyrtus plautus*,

27 genes (15 tRNAs, 2 rRNAs, and 10 PCGs) were encoded by the majority-strand (J-strand), and 10 genes (3 PCGs and 7 tRNAs) were encoded by the minority-strand (N-strand). The values of AT skew and GC skew are often used to reveal the nucleotide composition of the mitochondrial genome (Alexandre et al. 2005). In this study, the nucleotide compositions of eight complete or nearly complete mitogenomes in Encyrtidae were investigated by calculating the percentages of AT-skew and GC-skew (Fig. 3). The results of the nucleotide skew statistics showed that the AT-skews in PCGs, tRNAs, and rRNAs of encyrtid mitogenomes were almost all positive, while the GC-skews were almost all obviously negative.

The nucleotide composition *O. plautus* was significantly biased toward adenine and thymine, with an A + T content of 84.6% (Table 2). The PCGs on the N strand had a T skew and slight G skew. However, the PCGs on the J strand exhibited a C skew. The tRNAs was negative for AT skews, besides the tRNAs on the N strand, which had an equal G and C distribution. Two rRNA genes (lrRNA and srRNA) were encoded on the J strand and exhibit an A skew and slight C skew.

Intergenic spacers and overlapping genes are very common in arthropod mitochondrial genomes (Ma et al. 2015; Chen et al. 2018). In the mitochondrial genome of *O. plautus*, a total of 166 bp of intergenic spacers ranging from 1 to 47 bp were found in eight locations. The minimum intergenic spacers (1 bp) located at *trnT-trnP*, *trnQ-trnW*, *nad6-cob*,

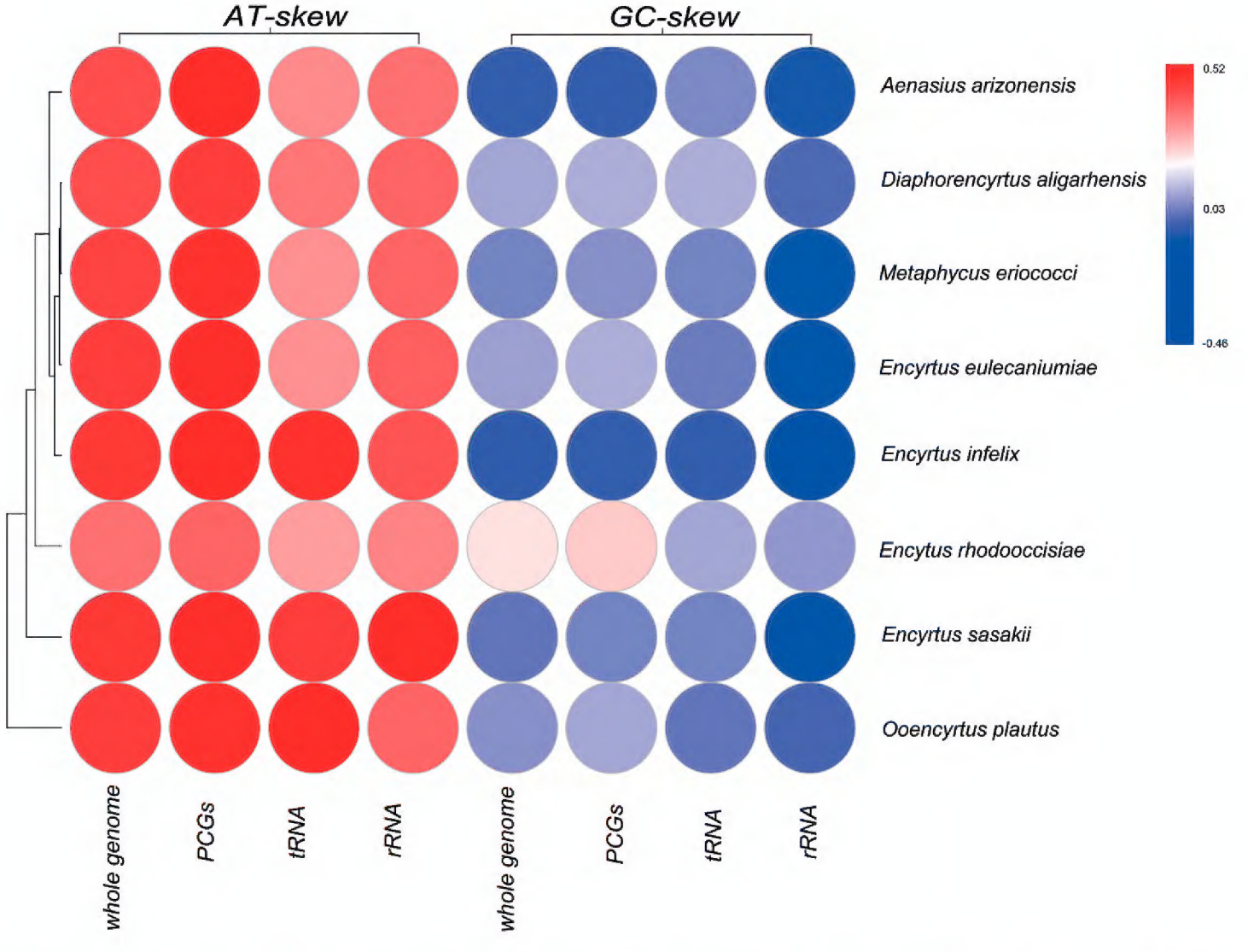

Figure 3. Nucleotide composition of various data sets of mitogenomes. Hierarchical clustering of Encyrtidae species based on the AT-skew and GC-skew.

Table 2. Nucleotide composition and skewing of the *Ooencyrtus plautus* mitogenome.

	size	T	C	A	G	AT-skew	GC-skew
whole genome	15730	0.401	0.085	0.445	0.069	0.052	-0.104
PCGs	11072	0.381	0.097	0.439	0.083	0.071	-0.078
PCGs(J)	8402	0.345	0.086	0.467	0.102	0.150	0.085
PCGs(N)	2670	0.494	0.081	0.352	0.073	-0.168	-0.052
tRNA	1467	0.455	0.066	0.437	0.042	-0.020	-0.222
tRNA(J)	994	0.454	0.036	0.437	0.073	-0.019	0.339
tRNA(N)	473	0.457	0.053	0.438	0.053	-0.021	0.000
rRNA	2103	0.437	0.068	0.453	0.042	0.018	-0.236

and *trnS2-nd1*. The maximum (47 bp) gene spacer was located at *trnW-trnC*. There were 18 overlapping gene regions, ranging from 2 bp to 7 bp in length in the nine mitogenomes. The longest overlapping sequence (7 bp) was between *atp6-atp8* and *nad4-nad4l*.

Protein-coding genes (PCGs)

In the mitochondrial genome of *O. plautus*, the total length of protein-coding genes was 11,072 bp, accounting for 70.39% of the entire genome. Most of them were encoded on the J strand, and only *nad2*, *nad6*, and *cob* were encoded on the N strand. The average A + T content of the 13 protein-coding genes was 82%, ranging from 73.7% (*cox1*) to 93.80% (*atp8*) for individual genes.

To investigate further this high A and T content, and the frequency of synonymous codon usage, we calculated relative synonymous codon usage (RSCU) values. The relative synonymous codon usages (RSCU) of the *O. plautus* are shown in Fig. 4. Taken together, the most frequently used codons are UUA (Leu2), AGA (Ser1), and CGA (Arg), whereas those ending in G or C, CUG, CUC, CAG, and GGC were the less frequently used codons. The codons ending with A or T are predominant, which at least partly leads to the bias towards A and T.

The predicted initiation codons are ATN as in most other insect mitochondrial genomes (Dowton and Austin 1999). There were six genes (*cox1*, *cob*, *nad4*, *nad6*, *atp6*, *and cox3*) starting with ATG and seven genes (*nad1*, *nad3*, *nad6*, *nad4L*, *nad5*, *cox2*, *and atp8*) starting with ATT. All protein-coding genes terminated at the most common stop codon TAA, except for *nad1* and *nad5*, which stopped with single T (Table 3).

Previous work (Zhang et al. 2020) showed that Eulophidae and Encyrtidae have close relationships, so we chose Eulophidae species as the reference sequence. Using *Chouioia cunea* as a reference sequence, we calculated the non-synonymous substitution rate (Ka), synonymous substitution (Ks), and Ka/Ks ratio of each PCG for *O. plautus* (Fig. 5). Ka and Ks are non-synonymous and synonymous substitution rates, respectively. They are controlled by functionally related sequence contexts, such as encoding amino acids and participating in exon splicing (Parmley and Hurst 2007). The ratio of the two parameters Ka/Ks (a measure of the strength of selection) is defined as

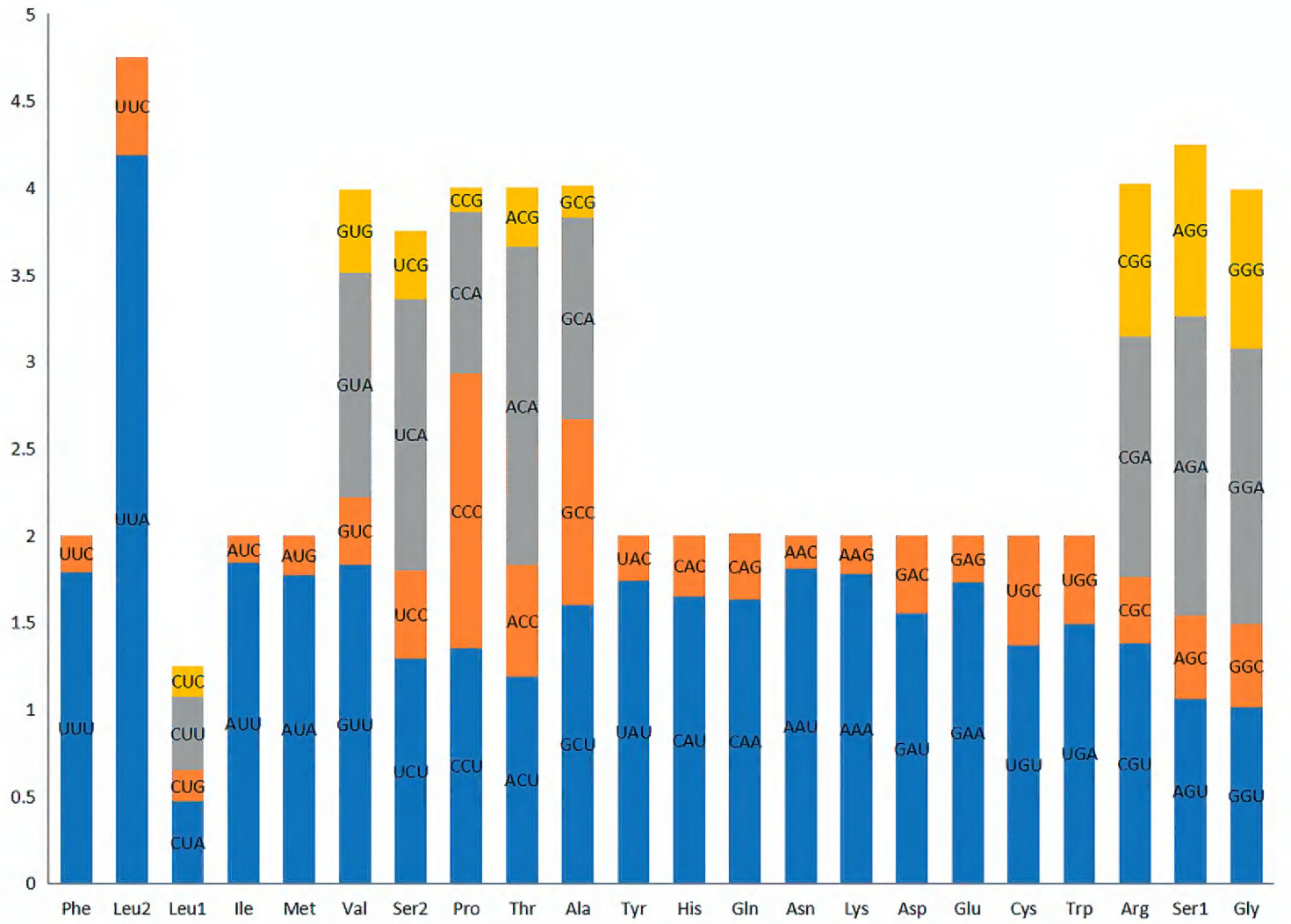

Figure 4. Relative synonymous codon usage (RSCU) of the mitochondrial genomes of *Ooencyrtus plautus*.

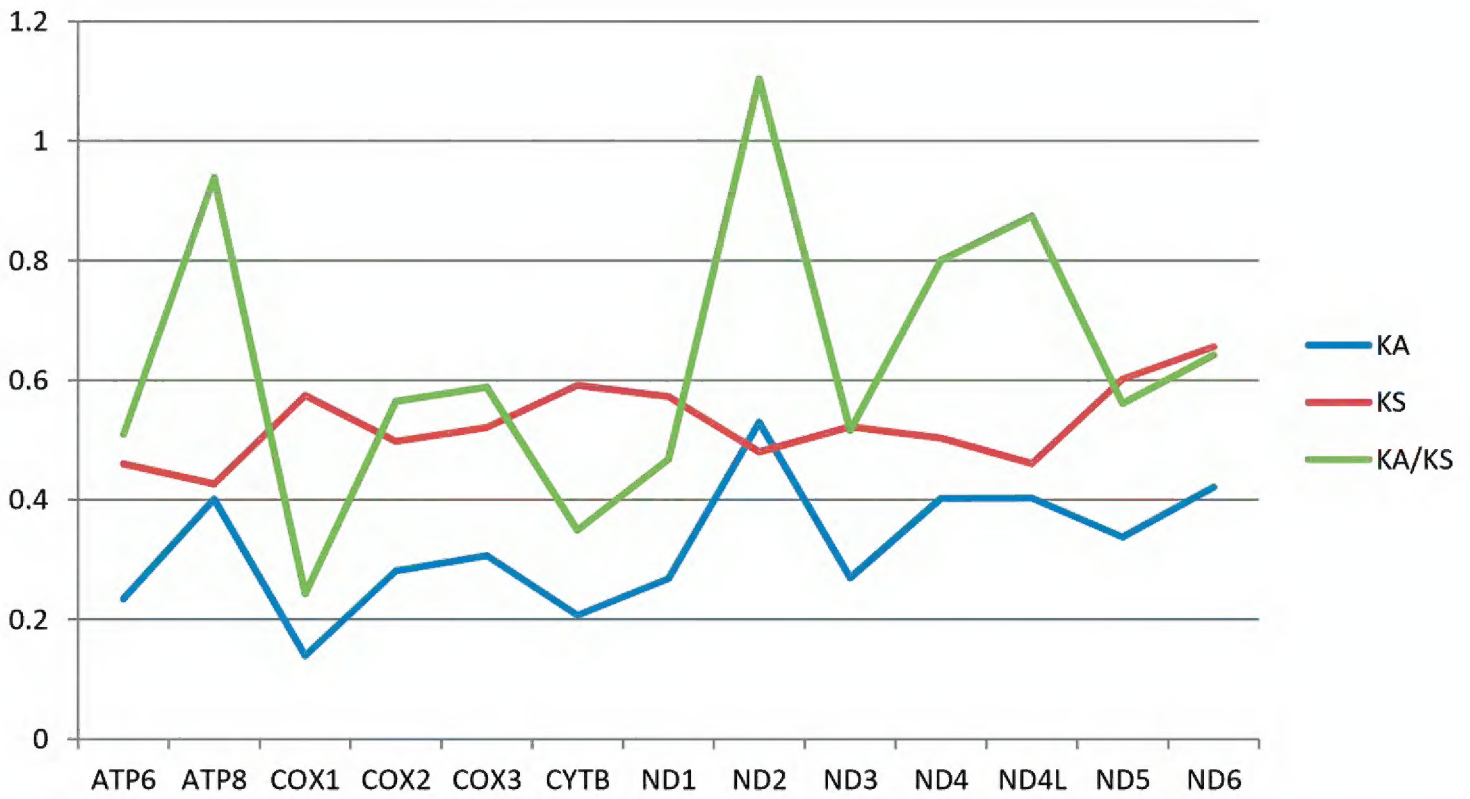

Figure 5. Evolutionary rates of mitochondrial genomes in *Ooencyrtus plautus*. The numbers of nonsynonymous substitutions per nonsynonymous site (Ka), the number of substitutions per synonymous site (Ks), and the ratio of Ka/Ks for each gene are given, using *Chouioia cunea* as the reference sequence.

the degree of evolutionary change (Wang et al. 2011). A value of Ka/Ks greater than 1 means positive selection exists, indicating that non-synonymous mutations are more favored by Darwinian selection, and they will be retained at a rate greater than synonymous mutations. In *O. plautus*, the Ka/Ks ratio of *nad2* was greater than 1, indicating that the *nad2* gene had a positive selection. This phenomenon was first discovered in Encyrtidae. In addition, the Ka/Ks of *atp8* is as high as 0.9389, which is the highest except for *nad2*. The high Ka/Ks phenomenon of *nad2* and *atp8* is also found in other species (Jia et al. 2020; Guo et al. 2021; Xu et al. 2021). The reason for this phenomenon may be that the evolution speed of a gene is related to its function (Wang et al. 2011).

Table 3. Mitogenomic organization of *Ooencyrtus plautus*.

Name	Start	Stop	Strand	Length	Codons
trnI	1	68	+	68	
nad2	1064	75	-	870	ATT/TAA
trnR	1146	1082	-	65	
trnQ	1227	1160	-	68	
trnN	1229	1296	+	68	
trnW	1299	1367	+	69	
trnC	1481	1415	-	67	
trnS1	1482	1540	+	59	
trnY	1548	1613	+	66	
nad3	1613	1966	+	333	ATT/TAA
trnG	1975	2039	+	65	
cox3	2043	2828	+	747	ATG/TAA
atp6	2828	3502	+	666	ATG/TAA
atp8	3496	3657	+	159	ATT/TAA
trnD	3658	3721	+	64	
trnK	3796	3727	-	70	
cox2	3800	4474	+	639	ATT/TAA
trnL2	4486	4549	+	64	
cox1	4561	6090	+	1509	ATG/TAA
trnE	6161	6099	-	63	
trnF	6162	6229	+	68	
nad5	6230	7898	+	1419	ATT/T
trnH	7899	7966	+	68	
nad4	7970	9307	+	1281	ATG/TAA
nad4l	9301	9588	+	255	ATT/TAA
trnT	9589	9659	-	71	
trnP	9661	9725	+	65	
nad6	10289	9744	-	501	ATG/TAA
cob	11424	10291	-	1098	ATG/TAA
trnS2	11491	11423	-	69	
nad1	11493	12417	+	906	ATT/T
trnL1	12418	12483	+	66	
rrnL	12484	13791	+	1308	
trnA	13792	13860	+	69	
rrnS	13861	14655	+	795	
trnV	14656	14725	+	70	
trnM	14724	14788	+	65	

Transfer and ribosomal RNA genes

In total, 22 transfer RNA genes were found, ranging in size from 59 bp (*trnS1*) to 71 bp (*trnT*). Most of the tRNAs were encoded on the J strand and only 7 tRNAs (*trnT, trnE, trnK, trnC, trnQ, trnR, and trnS2*) were encoded on the N strand. The average nucleotide composition of these tRNAs was A: 43.7%, T: 45.5%, C: 6.6%, and G: 4.2%, with a total average A + T content of 89.2%. All tRNA sequences can be folded into the canonical cloverleaf secondary structure, except for *trnS1* which lacked the dihydrouridine (DHU) arm. A lack of the DHU arm in *trnS1* was found in the mitochondrial genomes of most insects (Dowton et al. 2002b). Changes in the length of the DHU and TΨC arms lead to differences in the size of the tRNA sequence (Shao et al. 2001). In addition, some mismatches (G-U in *trnD*, *trnA*, and *trnV*, U-U in *trnS1*, and two U-U in *trnA*) were found in *O. plautus*.

The two rRNA genes, the larger ribosomal gene (rrnL) and the smaller ribosomal gene (rrnS), were located between *trnA* and *trnL1*, and *trnV* and *trnA*, respectively. The average of the total size of two rRNAs was 2,103 bp and the average A + T content was 89%.

Gene arrangement patterns

Gene rearrangement in the mitochondrial genome is a relatively rare event in the evolutionary history of insects (Cameron 2014). However, in the Hymenoptera lineage, more and more species have mitochondrial gene rearrangements (Shao et al. 2003). Gene rearrangements are a common phenomenon found in almost all the sequenced Chalcidoidea mt genomes (Fig. 2). In the Chalcidoidea mitochondrial genomes studies, most of the tRNA genes are rearranged (Tang et al. 2018). In the Hymenoptera, numerous rearrangements of protein-coding genes have been identified in several groups (Xiao et al. 2011; Wei et al. 2014). Compared with the putative ancestral pattern of the insect mitochondrial genome, dramatic gene rearrangements, not only in tRNA genes but also in protein-coding genes, were found in *O. plautus* mitochondrial genomes. In the mitogenome of *O. plautus*, a total of 17 genes, including six PCGs, 10 tRNAs, and one rRNA gene are rearranged (Fig. 7).

In Encyrtidae, the gene order of all species of *Encytus* are identical. We used their gene order as a template to analyze the gene rearrangement of Encyrtidae. It can be summarized into four obvious rearrangements. They are the rearrangement of gene clusters *trnI-nad2-trnW*, *trnY-trnS1-trnN-trnC-trnR*, and *trnA-trnQ-rrns-trnV-trnM*. In addition, there are inversions of *trnE-trnF* and *trnP-trnT*.

Firstly, the *trnW-trnI-nad2* gene cluster in these four species (*Aenasius arizonensis*, *Diaphorencyrtus aligarhensis*, *Metaphycus eriococci*, and *Platencytus parkeri*) was rearranged as *trnI-nad2-trnW*, most species in Chalcidoidea of which were in this order. However, in *O. plautus*, *trnW* was translocated, causing the gene cluster *trnI-nad2-trnW* to be divided into *trnI-nad2* and *trnW*, which is the first of its kind to be reported in Chalcidoidea. Besides, the *trnI-nad2* gene cluster was translocated to upstream. Secondly, the

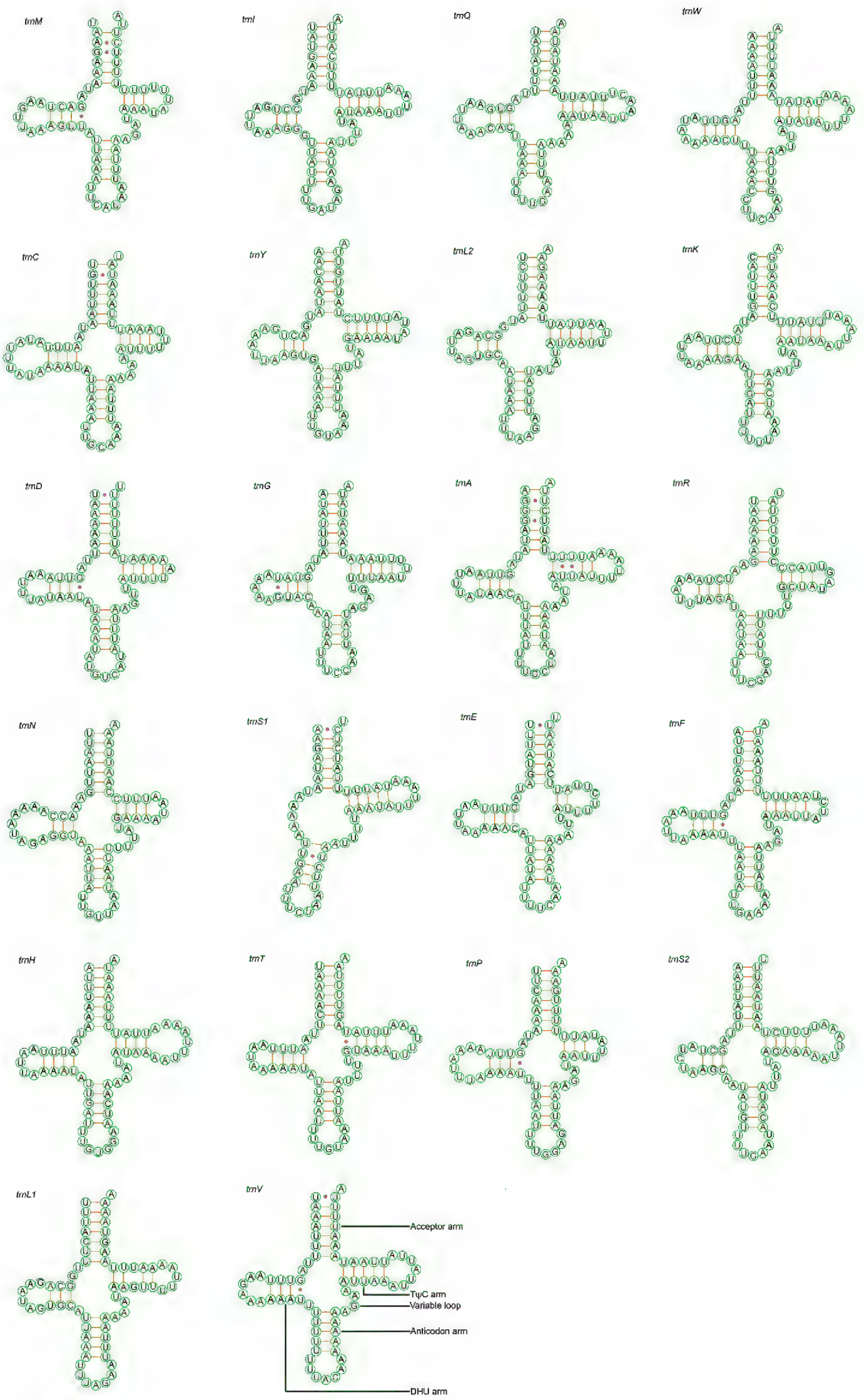

Figure 6. Predicted secondary cloverleaf structure for the tRNAs of *Ooencyrtus plautus*.

Figure 7. Mitochondrial genome organization in Encyrtidae. Colored boxes indicate rearranged gene clusters, and gray boxes indicate conserved gene blocks.

trnN-trnS1-trnY-trnC-trnR gene clusters in four species (*A. arizonensis*, *D. aligarhensis*, *M. eriococci*, and *O. plautus*) have undergone disorderly rearrangements. Thirdly, *trnE-trnF* inversion occurred in both *A. arizonensis* and *D. aligarhensis*. In addition, *trnP-trnT* inversion also occurred in *D. aligarhensis*. Finally, in *D. aligarhensis* and *O. plautus*, *trnQ* in the gene cluster *trnA-trnQ-rrnS-trnV-trnM* has a long-distance transposition. In *M. eriococci*, *trnM* and *trnQ* were transposed in the gene cluster *trnA-trnQ-rrnS-trnV-trnM*.

Phylogenetic analyses

Phylogenetic relationships were analyzed using the concatenated nucleotides and amino acids sequences of 13 PCGs from eight encyrtid species and two outgroups. Four topologies were constructed using both ML and BI methods and two different data sets.

The topological structures of these two phylogenetic trees reconstructed by ML analysis were identical with the phylogenetic tree reconstructed by BI based on PCG (Fig. 8). While Bayesian tree reconstructed on the AA data set showed different phylogenetic relationships in the clade consist of *Encyrtus rhodococcusiae*, *E. eulecaniumiae* and *E. sasakii* which was *E. eulecaniumiae* + (*E. rhodococcusiae* + *E. sasakii*) rather than *E. rhodococcusiae* + (*E. eulecaniumiae* + *E. sasakii*). The deviation and rate heterogeneity of the nucleotide composition of the mitogenomes are the basis of the rapid evolution of Chalcidoidea, which leads to inconsistent topologies based on different data types and analysis strategies (Wu et al. 2020). Currently, the CAT + GTR model implemented in PhyloBayes software was found to be the best fitting model for all data sets. In addition, the AA data set is considered as the best fit matrix for reconstructing phylogenetic trees (Li et al. 2015, 2017).

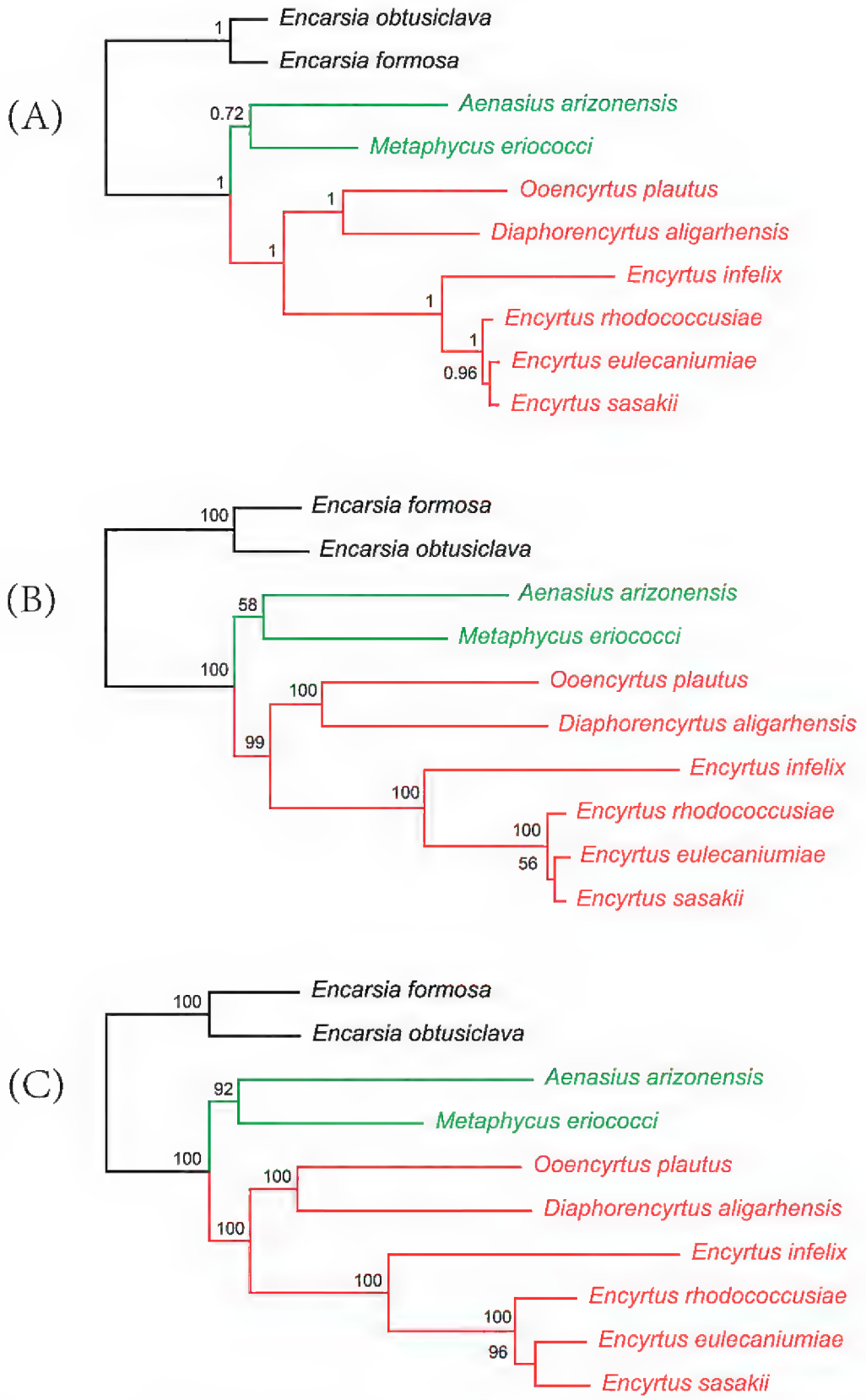

Figure 8. Phylogenetic tree produced by maximum likelihood and Bayesian inference analyses based on the 13 PCG and the AA data set **A** phylogenetic tree produced by Bayesian inference analyses based on 13 PCG data set **B** phylogenetic tree produced by maximum likelihood analyses based on the AA data set **C** phylogenetic tree produced by maximum likelihood analyses based on the 13 PCG data set.

In the BI topology based on the AA data set, *A. arizonensis* and *M. eriococci* formed a clade representing Tetracneminae. The remaining six species formed a monophyletic clade representing Encyrtinae. In Encyrtinae, *Encyrtus* formed a monophyletic clade as a sister group to the clade formed by *O. plautus* and *D. aligarhensis*. *Encyrtus sasakii* and *E. rhodooccisiae* are most closely related in this monophyletic clade. In the BI topology based on the PCG data set, *E. eulecaniumiae* and *E. sasakii* form a sister group then to *E. rhodooccisiae*. *Encyrtus infelix* was the first diverged in *Encyrtus*.

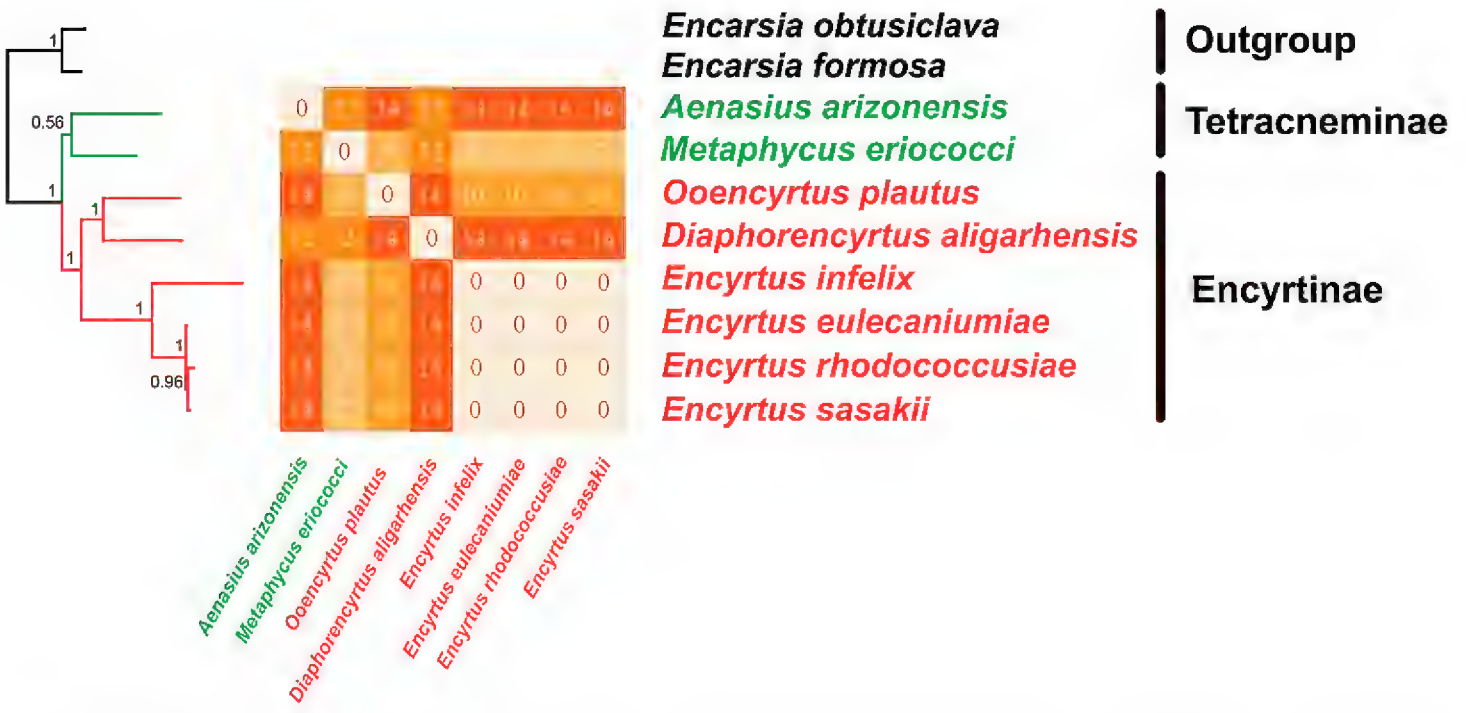

Figure 9. Mitogenomic phylogeny of eight Encyrtidae species and two outgroups based on the AA data set using the Bayesian inference. Heatmap of PBD values among encyrtid species is nested within the phylogeny, and the PBD values are listed.

Pairwise breakpoint distances (PBD) between the mitochondrial genomes of each species in Encyrtidae were calculated using the web server CREx, and heatmaps were constructed (Fig. 9). In the BI topology based on the AA data set, *E. infelix* was the first divergence branch in *Encytus*, which is consistent with the structure of the rest phylogenetic trees. The value of pairwise breakpoint distances between *E. infelix* and *D. aligarhensis* was 14, and the value of *E. infelix* and *O. plautus* was 10. Lower values of PBD indicated closer relationships which was consistent with the topology among *E. infelix*, *O. plautus*, and *D. aligarhensis* on the phylogenetic tree.

Discussion

In Encyrtidae, the gene rearrangement processes can be summarized into four rearrangement mechanisms, which can also be applied to Chalcidoidea (Fig. 2). First, the rearrangement between *trnI-nad2-trnW* includes the inversion and translocation. The gene cluster *trnI-nad2-trnW* presents three gene orders, one is *trnI-nad2-trnW*, another is *trnW-trnI-nad2*, and the last is one of them has a translocation. For example, in Pteromalidae, *nad2* in *Sycobia* sp. was translocated, but in other species, the order of *trnI-nad2-trnW* has not changed. The second is the rearrangement in gene cluster *trnN-trnS1-trnY-trnC-trnR* where gene rearrangements occurred most frequently. The main rearrangements were inversions and long-distance translocations (mainly *trnR*). In *Encyrtus*, translocation of *trnR* were occurred in all these four species, and this phenomenon was also found in Agaonidae. The final mechanism is the inversion of *trnE-trnF* and *trnP-trnT*. The *trnE-trnF* inversion occurred in *A. arizonensis*, *D. aligarhensis*, *Tamarixia radiata*, and *Aisopteromalus clandrae*. And the *trnP-trnT* inversion occurred in *Ceratosolen fusiceps*, *Eupelmus* sp. and *A. arizonensis*.

These four gene rearrangements can be applied to most of Chalcidoidea. When these tRNA gene rearrangement patterns were mapped on the estimated phylogenetic tree, the gene order of mitochondrial genome may resolve some contentious evolutionary questions (Shen et al. 2019). Based on the gene rearrangements occurring in Encyrtidae and the pairwise breakpoint distances heatmap, the hypotheses can be verified: the closely related species tend to have similar mitochondrial gene orders (Wu et al. 2020). For example, the gene order of all species of *Encytus* are identical. Correspondingly, the value of pairwise breakpoint distances among them are zero. Besides, the long-distance translocation of *trnQ* occurred in *D. aligarhensis* and *O. plautus*. In the topology of the BI tree based on AA matrix, *D. aligarhensis* and *O. plautus* are closely related and sister to the clade formed by genus *Encytus*. The PBD value between *O. plautus* and *Encyrtus* is 10, and the PBD value between *D. aligarhensis* and *Encyrtus* is 14. Lower PBD values indicates that *O. plautus* is more closely related to *Encyrtus*, which is in accordance with the phylogenetic tree. This phenomenon verified that PBD values could be used for inferring phylogenetic relationships. In summary, the gene rearrangement of mitogenome can provide a valuable source of characteristics for the reconstruction of molecular phylogeny (Wu et al. 2020).

As the sampling diversity of Encyrtidae is limited, it cannot completely solve the main classification question. Considering the limited research on taxonomic relationships inference based on molecular data, a comprehensive comparison of species morphology and genetic characteristics is needed to better understand the phylogenetic relationship of Encyrtidae. If more mitochondrial genomes are sequenced, the accuracy of phylogenetic relationships could be improved. We hope that future studies will combine morphology with more data sets from the mitochondrial genome to provide sufficient evidence for the phylogenetic relationship of the Encyrtidae.

Conclusions

In this study, we sequenced the nearly complete mitogenome of *O. plautus* that contains 37 genes and one control region. The nucleotide composition, codon usage, RNA structures, and protein-coding genes evolution were analyzed. The mitogenome genome of *O. plautus* reveals the phylogenetic relationship of Encyrtidae for the first time. In addition, the regularity of gene rearrangement within Encyrtidae is discussed. The results of this study will contribute to further studies on evolutionary relationships within Encyrtidae.

Funding

This work was supported by the National Natural Science Foundation of China (grant no. 32100352, 32100355), the Natural Science Fund of Anhui Province (grant no. 1908085QC93), the Natural Science Foundation of Universities of Anhui Province (grant no. KJ2020A0094), the Major Science and Technology Projects in Anhui Prov-

ince (grant no. 202003a06020009), the National Science and Technology Fundamental Resources Investigation Program of China (grant no. 2019FY101800) and the Natural Science Foundation of Anhui Normal University (grant no. 2020XJ19).

References

Alexandre H, Nelly L, Jean D (2005) Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of Metazoa, and consequences for phylogenetic inferences. Systematic Biology 54(2): 277–298. https://doi.org/10.1080/10635150590947843

Austin AD, Gibson GAP, Harvey MS (1998) Synopsis of Australian *Calymmochilus* Masi (Hymenoptera: Eupelmidae), description of a new Western Australian species associated with a pseudoscorpion, and review of pseudoscorpion parasites. Journal of Natural History 32(3): 329–350. https://doi.org/10.1080/00222939800770171

Andrews S (2013) FASTQC. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [accessed on 10 February 2022]

Bernt M, Merkle D, Ramsch K, Fritzsch G, Perseke M, Bernhard D, Schlegel M, Stadler PF, Middendorf M (2007) CREx: Inferring genomic rearrangements based on common intervals. Bioinformatics 23(21): 2957–2958. https://doi.org/10.1093/bioinformatics/btm468

Boore JL (1999) Animal mitochondrial genomes. Nucleic Acids Research 27(8): 1767–1780. https://doi.org/10.1093/nar/27.8.1767

Boore JL, Brown WM (1998) Big trees from little genomes: Mitochondrial gene order as a phylogenetic tool. Current Opinion in Genetics & Development 8(6): 668–674. https://doi.org/10.1016/S0959-437X(98)80035-X

Boore JL, Lavrov DV, Brown WM (1998) Gene translocation links insects and crustaceans. Nature 392(6677): 667–668. https://doi.org/10.1038/33577

Cameron SL (2014) Insect mitochondrial genomics: implications for evolution and phylogeny. Annual Review of Entomology 59: 95–117. https://10.1146/annurev-ento-011613-162007

Chen L, Chen PY, Xue XF, Hua HQ, Li YX, Zhang F, Wei SJ (2018) Extensive gene rearrangements in the mitochondrial genomes of two egg parasitoids, *Trichogramma japonicum* and *Trichogramma ostriniae* (Hymenoptera: Chalcidoidea: Trichogrammatidae). Scientific Reports 8(1): e7034. https://doi.org/10.1038/s41598-018-25338-3

Curole JP, Kocher TD (1999) Mitogenomics: Digging deeper with complete mitochondrial genomes. Trends in Ecology & Evolution 14(10): 394–398. https://doi.org/10.1016/S0169-5347(99)01660-2

Dierckxsens N, Mardulyn P, Smits G (2017) NOVOPlasty: De novo assembly of organelle genomes from whole genome data. Nucleic Acids Research 45(4): e18. https://doi.org/10.1093/nar/gkw955

Dowton M, Austin AD (1995) Increased genetic diversity in mitochondrial genes is correlated with the evolution of parasitism in the hymenoptera. Journal of Molecular Evolution 41(6): 958–965. https://doi.org/10.1007/BF00173176

Dowton M, Austin AD (1999) Evolutionary dynamics of a mitochondrial rearrangement "hot spot" in the hymenoptera. Molecular Biology and Evolution 16(2): 298–309. https://doi.org/10.1093/oxfordjournals.molbev.a026111

Dowton M, Belshaw R, Austin AD, Quicke DLJ (2002a) Simultaneous molecular and morphological analysis of Braconid relationships (Insecta: Hymenoptera: Braconidae) indicates independent mt-tRNA gene inversions within a single wasp family. Journal of Molecular Evolution 54(2): 210–226. https://doi.org/10.1007/s00239-001-0003-3

Dowton M, Castro LR, Austin AD (2002b) Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: The examination of genome 'morphology'. Invertebrate Systematics 16(3): 345–356. https://doi.org/10.1071/IS02003

Dowton M, Cameron SL, Dowavic JI, Austin AD, Whiting MF (2009) Characterization of 67 mitochondrial tRNA gene rearrangements in the Hymenoptera suggests that mitochondrial tRNA gene position is selectively neutral. Molecular Biology and Evolution 26(7): 1607–1617. https://doi.org/10.1093/molbev/msp072

Du Y, Song X, Liu X, Zhong BJ (2019) Mitochondrial genome of *Diaphorencyrtus aligarhensis* (Hymenoptera: Chalcidoidea: Encyrtidae) and phylogenetic analysis. Mitochondrial DNA Part B 4(2): 3190–3191. https://doi.org/10.1080/23802359.2019.1667913

Gissi C, Iannelli F, Pesole G (2008) Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. Heredity 101(4): 301–320. https://doi.org/10.1038/hdy.2008.62

Guo J, Yan ZT, Fu WB, Yuan H, Li XD, Chen B (2021) Complete mitogenomes of *Anopheles peditaeniatus* and *Anopheles nitidus* and phylogenetic relationships within the genus *Anopheles* inferred from mitogenomes. Parasites and Vectors 14(1): 452. https://doi.org/10.1186/s13071-021-04963-4

Heraty J (2017) Parasitoid biodiversity and insect pest management. Insect Biodiversity: 603–625. https://doi.org/10.1002/9781118945568.ch19

Heraty JM, Burks RA, Cruaud A, Gibson GAP, Liljeblad J, Munro J, Rasplus JY, Delvare G, Jansta P, Gumovsky A, Huber J, Woolley JB, Krogmann L, Heydon S, Polaszek A, Schmidt S, Darling DC, Gates MW, Mottern J, Murray E, Dal Molin A, Triapitsyn S, Baur H, Pinto JD, van Noort S, George J, Yoder M (2013) A phylogenetic analysis of the megadiverse Chalcidoidea (Hymenoptera). Cladistics 29(5): 466–542. https://doi.org/10.1111/cla.12006

Huelsenbeck JP (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61(3): 539–542. https://doi.org/10.1093/sysbio/sys029

Jia CH, Zhang XM, Xu SY, Yang TY, Gao TX (2020) Comparative analysis of the complete mitochondrial genomes of three rockfishes (Scorpaeniformes, Sebastiscus) and insights into the phylogenetic relationships of Sebastidae. Bioscience Reports 40(12): BSR20203379. https://doi.org/10.1042/BSR20203379

Johnson KP, Yoshizawa K, Smith VS (2004) Multiple origins of parasitism in lice. Proceedings Biological Sciences 271(1550): 1771–1776. https://doi.org/10.1098/rspb.2004.2798

Kazutaka K, Standley DMJB (2016) A simple method to control over-alignment in the MAFFT multiple sequence alignment program. Bioinformatics 32(13): 1933–1942. https://doi.org/10.1093/bioinformatics/btw108

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28(12): 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Li H, Shao R, Song N, Song F, Jiang P, Li Z, Cai W (2015) Higher-level phylogeny of paraneopteran insects inferred from mitochondrial genome sequences. Scientific Reports 5(1): e8527. https://doi.org/10.1038/srep08527

Li Q, Wei SJ, Tang P, Wu Q, Shi M, Sharkey MJ, Chen XX (2016) Multiple lines of evidence from mitochondrial genomes resolve phylogenetic relationships of parasitic wasps in Braconidae. Genome Biology and Evolution 8(9): 2651–2662. https://doi.org/10.1093/gbe/evw184

Li H, Leavengood JM, Chapman EG, Burkhardt D, Song F, Jiang P, Liu J, Zhou X, Cai W (2017) Mitochondrial phylogenomics of Hemiptera reveals adaptive innovations driving the diversification of true bugs. Proceedings of the Royal Society B, Biological Sciences 284(1862): 20171223. https://doi.org/10.1098/rspb.2017.1223

Liu HR, Li H, Song F, Gu WY, Feng JN, Cai WZ, Shao RF (2017) Novel insights into mitochondrial gene rearrangement in thrips (Insecta: Thysanoptera) from the grass thrips, *Anaphothrips obscurus*. Scientific Reports 7(1): e4284. https://doi.org/10.1038/s41598-017-04617-5

Ma H, Ma C, Li C, Lu J, Xia X (2015) First mitochondrial genome for the red crab (*Charybdis feriata*) with implication of phylogenomics and population genetics. Scientific Reports 5(1): e11524. https://doi.org/10.1038/srep11524

Ma Y, Zheng BY, Zhu JC, Tang P, Chen XX (2019) The mitochondrial genome of *Aenasius arizonensis* (Hymenoptera: Encyrtidae) with novel gene order. Mitochondrial DNA Part B 4(1): 2023–2024. https://doi.org/10.1080/23802359.2019.1617052

Mao M, Gibson T, Dowton M (2014) Evolutionary dynamics of the mitochondrial genome in the Evaniomorpha (Hymenoptera)—A group with an intermediate rate of gene rearrangement. Genome Biology and Evolution 6(7): 1862–1874. https://doi.org/10.1093/gbe/evu145

Marc L, Oliver D, Sabine K, Ralph B (2013) OrganellarGenomeDRAW—A suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. Nucleic Acids Research 41(W1): W575–W581. https://doi.org/10.1093/nar/gkt289

Meng G, Li Y, Yang C, Liu S (2019) MitoZ: A toolkit for animal mitochondrial genome assembly, annotation and visualization. Nucleic Acids Research 47(11): e63. https://doi.org/10.1093/nar/gkz173

Minh BQ, Nguyen M, von Haeseler A (2013) Ultrafast approximation for phylogenetic bootstrap. Molecular Biology and Evolution 30(5): 1188–1195. https://doi.org/10.1093/molbev/mst024

Oliveira D, Raychoudhury R, Lavrov DV, Werren JH (2008) Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). Molecular Biology and Evolution 25(10): 2167–2180. https://doi.org/10.1093/molbev/msn159

Parmley JL, Hurst LD (2007) Exonic Splicing Regulatory Elements Skew Synonymous Codon Usage near Intron-exon Boundaries in Mammals. Molecular Biology Evolution (8): 1600–1603. https://doi.org/10.1093/molbev/msm104

Rozas J, Sánchezdelbarrio J, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics (Oxford, England) 19(18): 2496–2497. https://doi.org/10.1093/bioinformatics/btg359

Sankoff D, Leduc G, Antoine N, Paquin B, Lang BF, Cedergren R (1992) Gene order comparisons for phylogenetic inference: Evolution of the mitochondrial genome. Proceedings of the National Academy of Sciences of the United States of America 89(14): 6575–6579. https://doi.org/10.1073/pnas.89.14.6575

Shahjahan RM, Hughes KJ, Leopold RA, Devault JD (1995) Lower incubation temperature increases yield of insect genomic DNA isolated by the CTAB method. BioTechniques 19(3): 332–334. https://doi.org/10.1016/S0360-3016(00)00637-4

Shao RF, Barker SC (2003) The highly rearranged mitochondrial genome of the plague thrips, *Thrips imaginis* (Insecta: thysanoptera): convergence of two novel gene boundaries and an extraordinary arrangement of rRNA genes. Molecular Biology and Evolution 20(3): 362–370. https://doi.org/10.1093/molbev/msg045

Shao RF, Campbell NJH, Barker SC (2001) Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera). Molecular Biology and Evolution 18(5): 858–865. https://doi.org/10.1093/oxfordjournals.molbev.a003867

Shao RF, Dowton M, Murrell A, Barker SC (2003) Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of insects. Molecular Biology and Evolution 20(10): 1612–1619. https://doi.org/10.1093/molbev/msg176

Shen ZC, Chen L, Chen L, Li YX (2019) Information from the mitochondrial genomes of two egg parasitoids, *Gonatocerus* sp. and *Telenomus* sp., reveals a controversial phylogenetic relationship between Mymaridae and Scelionidae. Genomics 111(5): 1059–1065. https://doi.org/10.1016/j.ygeno.2018.06.009

Sudhir K, Glen S, Li M, Christina K, Koichiro T (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution 35(6): 1547–1549. https://doi.org/10.1093/molbev/msy096

Tang P, Zhu JC, Zheng BY, Wei SJ, Sharkey M, Chen XX, Vogler AP (2018) Mitochondrial phylogenomics of the Hymenoptera. Molecular Phylogenetics and Evolution 131: 8–18. https://doi.org/10.1016/j.ympev.2018.10.040

Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ (2016) W-IQ-TREE: A fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Research 44(1): 232–235. https://doi.org/10.1093/nar/gkw256

Wang DP, Fei L, Wang L, Huang S, Yu J (2011) Nonsynonymous substitution rate (Ka) is a relatively consistent parameter for defining fast-evolving and slow-evolving protein-coding genes. Biology Direct 6(1): e13. https://doi.org/10.1186/1745-6150-6-13

Wei SJ, Li Q, van Achterberg K, Chen XX (2014) Two mitochondrial genomes from the families Bethylidae and Mutillidae: Independent rearrangement of protein-coding genes and higher-level phylogeny of the Hymenoptera. Molecular Phylogenetics and Evolution 77: 1–10. https://doi.org/10.1016/j.ympev.2014.03.023

Wu YF, Yang HL, Feng ZB, Li YB, Zhou WB, Song F, Li H, Zhang LM, Cai WZ (2020) Novel gene rearrangement in the mitochondrial genome of *Pachyneuron aphidis* (Hymenoptera: Pteromalidae). International Journal of Biological Macromolecules 149: 1207–1212. https://doi.org/10.1016/j.ijbiomac.2020.01.308

Xiao JH, Jia JG, Murphy RW, Huang DW (2011) Rapid evolution of the mitochondrial genome in chalcidoid wasps (Hymenoptera: Chalcidoidea) driven by parasitic lifestyles. PLoS ONE 6(11): e26645. https://doi.org/10.1371/journal.pone.0026645

Xiong M, Zhou QS, Zhang YZ (2019) The complete mitochondrial genome of *Encyrtus infelix* (Hymenoptera: Encyrtidae). Mitochondrial DNA Part B 4(1): 114–115. https://doi.org/10.1080/23802359.2018.1537727

Xu W, Lin SP, Liu HY (2021) Mitochondrial genomes of five *Hyphessobrycon tetras* and their phylogenetic implications. Ecology and Evolution 11(18): 12754–12764. https://doi.org/10.1002/ece3.8019

Yuan W, Chen J, Jiang LY, Qiao GX (2015) Hemipteran mitochondrial genomes: Features, structures and implications for phylogeny. International Journal of Molecular Sciences 16(6): 12382–12404. https://doi.org/10.3390/ijms160612382

Zhang JX, Lindsey ARI, Peters RS, Heraty JM, Krogmann L (2020) Conflicting signal in transcriptomic markers leads to a poorly resolved backbone phylogeny of chalcidoid wasps. Systematic Entomology 45(4): 738–802. https://doi.org/10.1111/syen.12427

Zhu JC, Tang P, Zheng BY, Wu Q, Wei SJ, Chen XX (2018) The first two mitochondrial genomes of the family Aphelinidae with novel gene orders and phylogenetic implications. International Journal of Biological Macromolecules 118: 386–396. https://doi.org/10.1016/j.ijbiomac.2018.06.087